Agricultural Economists' Effectiveness in Reporting and Conveying Research Procedures and Results

Joe L. Parcell, Terry L. Kastens, Kevin C. Dhuyvetter, and Ted C. Schroeder

This study reviews articles using regression analysis published in the *Journal of Agricultural and Resource Economics* from 1994 to 1998 to determine agricultural economists' effectiveness in reporting and conveying research procedures and results. Based on the authors' experiences of surveying articles for this study, several suggestions for reporting of results and how to better separate statistical from economic significance are offered. First, clearly define the dependent variable—preferably in the results table as well as within the text. Second, report parameter estimates in an interpretable form either in the results table or in a subsequent table. Third, report summary statistics. Fourth, report degrees of freedom conspicuously in the results table. Fifth, report if statistically insignificant variables have been dropped. Lastly, weigh economic importance aside from statistical significance and use simulation to express economic significance where appropriate.

How effective are agricultural economists at reporting and conveying research procedures and results? Are procedures and related findings consistent with economic and statistical theory? Do agricultural economists reveal potential shortcomings in interpreting and conveying research results? To address these questions, a survey was conducted of articles published in the *Journal of Agricultural and Resource Economics* over the five years 1994–1998 that used regression analysis. This journal was chosen because of its emphasis on applied research.

The motivation for this study was an article by McCloskey and Ziliak, "The Standard Error of Regression," where the two authors surveyed a set of articles published in the *American Economic Review* to determine the extent to which researchers used and interpreted statistical and economic significance. McCloskey and Ziliak noted that a parameter estimate can be economically significant without being statistically significant and a parameter estimate can be statistically significant without being economically significant. They hypothesized that economists confuse tests of statistical significance and tests of substantive effects. Blaug, in reviewing McCloskey's (1985) claims about the confusion between significance tests and tests of substantive effects, challenged McCloskey's methodology. Blaug postulated, Who are you [McCloskey] to be passing judgment and should anyone be making such judgments? This question lingers in the present study; however, our intent is to report what is done and not to pass judgment. Pieces of the present study do elicit whether researchers differentiate between tests of statistical significance and tests of substantive effects. Also, new questions were designed to determine the ease with which the reader could synthesize procedures and results for application to other problems in the relevant subject area.[1]

Joe L. Parcell is an assistant professor with the Department of Agricultural Economics at the University of Missouri. Terry L. Kastens, Kevin C. Dhuyvetter, and Ted C. Schroeder are assistant, associate, and professor, respectively, with the Department of Agricultural Economics at Kansas State University. The authors would like to thank Harold Breimyer and Bill Tomek for comments on earlier drafts of this manuscript and participants of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management for helpful suggestions.

[1] These questions were posed with the belief that research conveys pieces of knowledge even if these pieces are part of a larger problem. We do not approach these questions with skepticism or cynicism as does Levins, who seems to question the reporting of pieces of a larger problem.

Agricultural and Resource Economics Review 29/2 (October 2000) 173–182

Statistical significance has long been an issue for agricultural economists. Agricultural economists, generally heavily reliant on sample data, were first adopters of econometric and statistical analyses to evaluate hypotheses. However, Tweeten argued against over-use of statistical tests in research, noting, ". . . to insist that all hypotheses be subject to rigorous statistical tests is to restrict economics to a narrow quantitative science." (p. 551)

The reader's ability to assess implications of research results is dependent upon the author's ability to explain procedures, extract generalities, and quantify effects. If the reader cannot clearly follow the author's outline, the value of the research is greatly reduced because overarching implications will likely be missed and replicating results will undoubtedly be more difficult. Tomek has suggested confirmation and replication of research is a quality-quantity trade-off. And, Ladd noted, "When I was a student, we were taught replication was a necessary process." (p. 8) Surely, making research easier to confirm and replicate improves quality, perhaps without loss of quantity.

This study examines the ease by which readers can synthesize procedures and results and differentiate statistical from economic significance. Additionally, by making researchers using regression analysis aware of their departures from statistical and economic interpretation in publishing research, perhaps individuals and the profession as a whole can take steps to improve reporting, thereby increasing the value of the research to interested parties.

The Survey Instrument

Questions posed in the current study were adapted from survey questions asked by McCloskey and Ziliak relative to articles using regression analysis published in the *American Economic Review* during the 1980s. McCloskey and Ziliak surveyed articles to determine if general economists confused tests of statistical significance and tests of substantive effects. Upon review of the questions posed of articles by McCloskey and Ziliak it was decided that several of the questions were not relevant in determining the ease by which readers could synthesize procedures and results. Therefore, the current study uses a subset of questions asked by McCloskey and Ziliak and expands the survey to ascertain methods of reporting procedures and results.

For the current study, 25 questions were asked regarding each journal article. Questions were stated such that the surveyor could respond with a "yes" representing the authors do this or report this, a "no" representing the authors do not do this or do not report this, and "not applicable" representing that the question asked was not applicable to the content in the article. Two of the questions addressed data used and estimation procedures and three questions were directed at determining the level of confidence indicated by the use of asterisks. Additionally, five of the questions addressed goodness of fit measures used by the author(s). The specific questions asked of each article were:

1. *What type of data is used for this analysis, time-series, cross-sectional, or panel?* This question was asked to determine if the methods of reporting results differed by data type.
2. *What type of regression technique was used, Ordinary (or transformation of) Least Square (OLS), Limited Dependent Variable (LDV), or System (SYS)?* How results are reported and their use in economic implications may differ by regression technique. The interpretation of coefficients estimated using a LDV model is less intuitive than for a model estimated using OLS. Thus, it may be that simulation is used more often with LDV models than OLS models.
3. *Are p-values of parameter estimates reported?* A p-value provides a more concise measure of statistical significance than the student t-statistic. For the t-statistic, degrees of freedom must be known to draw conclusions about statistical significance.
4. *Are standard errors reported?* Standard errors can be used in constructing confidence intervals around parameter estimates.
5. *Are t-statistics reported?* A t-statistic is a standardized measure of the level of statistical significance. Standard errors are computed by taking the ratio of the coefficient estimate to the t-statistic so that confidence intervals around the parameter estimate can be constructed.
6. *Are asterisks reported?* Assigning asterisks is a common indicator of statistical significance of parameter estimates.

7–9. *What level of confidence does* 1, 2, *or* 3 *asterisk(s) represent?* Choices were 80, 85, 90, 95, 99, and other? These ques-

tions were asked to determine if a standardized method of assigning asterisks was used.

10–14. *Was a goodness of fit measure reported in the results table*, i.e., R-squared, Adjusted *R*-squared, Root Mean Square Error (RMSE), Log-Likelihood, or Other? If the explanatory variables poorly explain variability in the dependent variable, results based on statistical significance used to make economic implications may have less credibility.

15. *Does the author explain the dependent variable?* The reader can only interpret regression results if it is clear what the regression model is explaining.

16. *Does the author state that he/she has dropped statistically insignificant variables from the reported models?* Data mining occurs, do authors make this apparent? How many times as a reader have you asked, why did the author(s) not include some variable? The variable may have been included during iterations of model estimation, and subsequently dropped if not statically significant. Stating that statistically insignificant variables were dropped is a dilemma for the author because it is admitting to pre-testing yet it can be beneficial for future research efforts.

17. *Are degrees of freedom or number of observations reported in the table?* Degrees of freedom indicates the power of hypotheses tests conducted on the model and parameter estimates. McCloskey and Ziliak noted that for the test of purchasing power parity, choosing a large enough N (observations) will cause the hypotheses of "β equal 1" to be rejected, even when the parameter estimate is close to one, say 0.999. On the other hand, small numbers of observations make one wonder about the robustness of the results. Conspicuously reporting degrees of freedom improves a reader's ability to interpret results. When degrees of freedom have to be searched for or somehow calculated by a reader the likelihood of making an error increases, and as writers we have increased the chances of misinterpretation of our results.

18. *Are summary statistics reported in the paper?* Summary statistics are important for the reader to measure the size of the economic effect if the author does not explicitly explain the effect.

19. *Are the coefficients reported in elasticity form, or in some other interpretable form, so that readers can discern the economic impact of regressors?* Readers must be able to interpret research results so they can evaluate the applicability of the research. The reader's interpretation is limited if coefficients are not reported in some interpretable form.

20. *Are both statistical and economic importance/significance of results discussed?* Reporting both statistical and economic significance helps the reader assess the impact of the study.

21. *In ranking numerical results, does the author emphasize statistical significance more than economic significance?* Because a coefficient is statistically significant does not mean the variable associated with the coefficient is necessarily economically significant. Likewise, a variable can be economically significant and its coefficient not statistically significant.

22. *Does the author discuss the size of the coefficients or size of the effects?* Relative to the problem being evaluated, does the author indicate whether the economic impact of the coefficient is large enough to matter.

23. *In the model results, does the author discuss expected, yet dismiss unexpected, signs on statistically insignificant parameter estimates?* That is, did the author emphasize coefficients that were of the theoretically correct sign and not significant and disregard coefficients that were of the theoretically wrong sign but not significant? This question examines whether the author was consistent in using statistical significance.

24. *Does the author avoid using the word "significance" in ambiguous ways, meaning "statistically significant" in one sentence and "large enough to matter for policy or science" in another?* Economic importance differs from statistical significance, and the ability of the author to clearly convey this enables the reader to easily evaluate the credibility of the study.

Table 1. Summary Statistics of Surveyed Articles Published in the *Journal of Agricultural and Resources Economics* between 1994 and 1998 and that used Regression Analysis

Variable	1994	1995	1996	1997	1998	All Years
Total articles published	32	26	30	27	36	151
Articles using regression	22	12	22	16	14	86
% of articles reviewed	68.8	46.2	73.3	59.3	38.9	57.0
Total pages published	463	410	418	426	578	2294
Pages reviewed						
Total	274	176	282	206	211	1149
Shortest paper	8	11	9	8	9	8
Longest paper	17	20	19	17	22	22
% of pages published	59.2	42.9	67.5	48.5	36.5	50.1

25. *Did the paper use simulation to enhance interpretation of important coefficients?* Simulation can enhance the interpretation of the results; however, as McCloskey and Ziliak noted, statistical significance should not be the sole criteria for inclusion of variables for simulation.

Survey Results

Summary statistics of articles surveyed that used regression analysis, and published in the *Journal of Agricultural and Resource Economics* between 1994 and 1998, are reported in table 1. Of 151 articles published, 86 (57%) used regression and were thus included in this survey. The five years of published articles totaled 2,294 pages. The number of published pages representing studies using regression analysis was 1,149, or 50% of the total pages published.

Some of the questions posed to the articles surveyed were to determine generalities about data used, econometric technique, and measures of goodness of fit and statistical significance. These questions are summarized in table 2. Data were classified as time-series, cross-sectional, or panel. Of the 86 articles surveyed, 35% used time-series data, 37% used cross-sectional data, and 29% used panel data. Nearly 50% of the articles surveyed used OLS (or derivations of single-equation OLS models, e.g., GLS) for estimating relationships between the dependent and independent variables. OLS was used for estimation in 33%, 46%, and 64% of the articles using time-series, cross-sectional, and panel data, respectively. The LDV model was used predominantly for estimation with cross-sectional data. SYS estimation was used primarily with time-series data (67%).

The next set of questions posed sought to determine the measures of goodness of fit and statistical significance used. Only 2.3% of the articles surveyed reported *p*-values. This result was somewhat surprising because *p*-values allow the reader to quickly asses the statistical significance of the coefficient; however, *p*-values do not allow for easy computation of confidence intervals as do *t*-statistics and standard errors. Of the articles surveyed, 34% and 64% reported standard errors and *t*-statistics, respectively. There was a clear preference for reporting *t*-statistics.

McCloskey and Ziliak argued that these measures of statistical significance, i.e., *t*-statistics, *p*-values, and standard errors are irrelevant in assessing the size effect, and they noted that for the *t*-statistic the *t* table does not indicate what is close to being a significant effect. Additionally, *t*-statistics, *p*-values, or standard errors may not be the appropriate measure of type I error if econometric assumptions are violated, e.g., residual normality, autocorrelation, collinearity.

Asterisks (*) highlighting levels of statistical significance were reported in 62% of the articles surveyed.[2] McCloskey and Ziliak reported only 25% of the articles they reviewed in the *American Economic Review* used asterisks to denote statistical significance. Authors were inconsistent in specifying the level of significance indicated by one, two, or three asterisks. For instance, Hurd reported that one asterisk (*) indicated statistical significance at the 0.05 level, and Prichett, Liu, and Kaiser reported that two asterisks (**) indicated statistical significance at the 0.05 level. However, the majority had one, two, and three asterisk(s) indicating significance at the 0.10, 0.05, and 0.01 level, respectively.

Some authors chose to convey significance levels of coefficients outside of the confidence inter-

[2] Articles (two) indicating level of significance using alphabetical notation were reported as "no."

Table 2. Responses of Data used, Econometric Technique, and Measure of Statistical Significance used in Surveyed Articles Published in the *Journal of Agricultural and Resources Economics* between 1994 and 1998 and that used Regression Analysis

Variable				Percentage of Articles[a]
1. Type of data used (%)[b]				
Time-series				34.9
Cross-sectional				37.2
Panel				29.1
2. Type of econometric technique used (%)[b]				
Ordinary Least Squares				47.7
Limited Dependent Variable				26.7
System of Equations				37.2
	Time-series		Cross sectional	Panel
Econometric technique by data type (%)[b]				
Ordinary Least Squares	33.3		46.6	64
Limited Dependent Variable	3.3		56.3	21.9
System of Equation	66.7		21.9	24
3. *p*-values reported (% yes)				2.3
4. Standard error reports (% yes)				34
5. t-statistics reported (% yes)				63.9
6. Asterisks reported (% yes)				61.6
7–9. The asterisk(s) represented significance at the stated significant level (%)[c]				
		Significance level		
	85%	95%	90%	99%
*	2	51	41.2	7.8
**	—	10.8	62.2	27
***	10	10	10	60
10–14. Goodness of fit measure used in the table (% yes)				77
R-squared				50.6
Adjusted R-squared				16.9
RMSE				11.7
Log-Likelihood				27.2
Other				22.1

[a]Based on 86 articles except for the Goodness of Fit question which applied to 77 articles.
[b]Percentages do not sum to 100% due to some articles using more than one type of data or econometric method.
[c]Of the paper surveyed, 53 (62.4%) used asterisks.

vals noted in the table. For example, Kenkel and Norris used one and two asterisks to denote significance at the 10% and 5% level in the results table, but in discussing their results, "The coefficients for crop acres and number of crops were significant at the 0.15 and 0.16 levels." (p. 367) Tables are reported to provide an overview summary. In discussing results it may be necessary to stray from the precise level of significance indicated by asterisks reported in the table to include variables that are economically significant. McCloskey and Ziliak used the test of purchasing power parity ($\beta = 1$) in paraphrasing Wald's notion that close depends on, ". . . the special purpose of the investigation—good enough for inflation control, say if $\beta = 0.85$, though not good enough to make money on the foreign exchange market unless $\beta = 0.99998$." (p. 98) Possibly, assigning asterisks distracts the reader from focusing on economic significance.

The most used goodness of fit measure was R-squared, with 51% of the articles reporting this measure, and Log-Likelihood the second most reported measure (27%). Somewhat surprisingly, 23% of the paper did not report any goodness of fit measure, providing readers no direct evidence of the model's explanatory power.

Table 3 summarizes survey questions posed of articles regarding ease of interpretation of results, consistencies in interpreting statistical hypotheses, and differentiation between tests of statistical significance and tests of substantive differences. Tables 4 and 5 separate results in table 3 by econometric method and data type, respectively. The author explained the dependent variable in 94% of the applicable articles (question 15), albeit some-

Table 3. Responses of Surveyed Articles Published in the *Journal of Agricultural and Resources Economics* between 1994 and 1998 and that used Regression Analysis

Question	# of articles where question applies	Percentage "yes"[a]
15. Does the author explain the dependent variable?	82	93.9
16. Does the author state that he/she has dropped statistically insignificant variables from reported models?	83	10.8
17. Is degrees of freedom or number of observations reported in the table?	85	32.9
18. Are summary statistics reported in a table?	86	52.3
19. Are the coefficients reported in elasticity form, or in some interpretable form, so that readers can discern the economic impact of regressors?	85	70.4
20. Are both statistical and economic importance/significance of results discussed?	86	73.3
21. In ranking numerical results the author(s) emphasized statistical significance more than economic significance?	85	35.3
22. Does the author discuss the size of the coefficients or size of the effects?	84	65.5
23. In the model results does the author discuss expected signs on statistically insignificant parameter estimates while dismissing unexpected signs on statistically insignificant parameter estimates?	84	14.3
24. Does the author avoid using the word "significance" in ambiguous ways, meaning "statistically significant" in one sentence and "large enough to matter for policy or science" in another?	86	80.2
25. Did the paper use simulation to enhance interpretation of important coefficients?	86	38.4

[a]Percentage "yes" of the number of articles where question applies.

times in obscure places in the text. In some cases searching for an explanation of the dependent variable, the fundamental component of the hypothesis test, detracted from the underlying theme of a paper. Studies expressing the dependent variable and its unit of measure in results tables were far easier and less apt to error in interpreting results. For example, Ramezani and Helmberger and Chen explained their dependent variables in the results table allowing readers to more easily interpret results.

Only 11% of the applicable articles indicated

Table 4. Responses, Separated by "estimation" Technique, of Surveyed Articles Published in the *Journal of Agricultural and Resources Economics* between 1994 and 1998 and that used Regression Analysis[a]

Question	OLS	LDV	System
15. Does the author explain the dependent variable?	92.3	95.2	82.8
16. Does the author state that he/she has dropped statistically insignificant variables from reported models?	10.3	14.3	6.9
17. Is degrees of freedom or number of observations reported in the table?	43.6	42.9	10.3
18. Are summary statistics reported in a table?	56.4	67	41.4
19. Are the coefficients reported in elasticity form, or in some interpretable form, so that readers can discern the economic impact of regressors?	69.2	38.1	86.2
20. Are both statistical and economic importance/significance of results discussed?	82.1	66.7	68.9
21. In ranking numerical results the author(s) emphasized statistical significance more than economic significance?	38.5	38.1	27.6
22. Does the author discuss the size of the coefficients or size of the effects?	64.1	42.9	75.9
23. In the model results does the author discuss expected signs on statistically insignificant parameter estimates while dismissing unexpected signs on statistically insignificant parameter estimates?	10.3	9.5	20.7
24. Does the author avoid using the word "significance" in ambiguous ways, meaning "statistically significant" in one sentence and "large enough to matter for policy or science" in another?	87.2	85.7	69
25. Did the paper use simulation to enhance interpretation of important coefficients?	41	47.6	38

[a]Multiple estimation techniques were used in some studies.

Table 5. Responses, Separated by Data Type, of Surveyed Articles Published in the *Journal of Agricultural and Resources Economics* between 1994 and 1998 and that used Regression Analysis[a]

Question	Time-series	Cross-sectional	Panel
15. Does the author explain the dependent variable?	90	93.8	84
16. Does the author state that he/she has dropped statistically insignificant variables from reported models?	10	12.5	8
17. Is degrees of freedom or number of observations reported in the table?	16.7	43.8	36
18. Are summary statistics reported in the paper?	30	72	52
19. Are the coefficients reported in elasticity form, or in some interpretable form, so that readers can discern the economic impact of regressors?	83.4	56.3	68
20. Are both statistical and economic importance/significance of results discussed?	80	65.6	76
21. In ranking numerical results the author(s) emphasized statistical significance more than economic significance?	20	37.5	48
22. Does the author discuss the size of the coefficients or size of the effects?	70	53.1	72
23. In the model results does the author discuss expected signs on statistically insignificant parameter estimates while dismissing unexpected signs on statistically insignificant parameter estimates?	16.7	15.7	8
24. Does the author avoid using the word "significance" in ambiguous ways, meaning "statistically significant" in one sentence and "large enough to matter for policy or science" in another?	76.7	78.1	88
25. Did the paper use simulation to enhance interpretation of important coefficients?	33.3	31.3	52

[a]Multiple data types were used in some studies.

that statistically insignificant variables were dropped from the reported models (question 16). Liu, Sun, and Kaiser reported they dropped a statistically insignificant variable with no significant impact on model results; therefore, they reported model results that included the insignificant variable.[3] Some authors may choose not to identify insignificant variables dropped from the model due to model fragility. Perhaps authors should list all variables that at one time were included in the model with model fragility test statistics. This would benefit readers who extend the study with the inclusion of variables not included in the original study. However, reporting that insignificant variables were dropped is admitting to pretesting which may be why most authors are reluctant to report this.

What use are measures of statistical significance when pretesting occurs? Wallace noted, in discussing the relevance of statistical significance following pretest estimation, that when pretesting occurs the true probability of type I error is unknown. Thus, a low p-value or large t-statistic may be preferred but be of little value. Wallace concluded,

> . . . in my opinion, statistical procedures substitute rather poorly for rigorous modeling based on the foundations of the field of application. Rather, statistical technique is a complement, not a substitute. . . . Occasional sinning [pretesting], therefore, may be inevitable but not necessarily fatal. (p. 443)

Possibly authors should rigorously report what they have done and leave the decision of the reliability of the results to the reader.

Less than half the articles surveyed (33%) reported degrees of freedom in the results table (table 3, question 17), and only 17% of the time-series studies reported degrees of freedom in the results table (table 5). Sample size indicates the power of the test. For example, Griliches stated

> Here and subsequently, all statements about statistical "significance" should not be taken literally. Besides the usual issues of data mining clouding their interpretation, the "sample" analyzed comes close to covering completely the relevant population. Tests of significance are used here as a metric for discussing the relative fit of different versions of the model. In each case, the actual magnitude of the estimated coefficients is of more interest than their precise "statistical significance." (p. 146, cited in McCloskey and Ziliak, p. 106)

Griliches noted that the sample size he used comes close to covering the relevant population; however, he does not indicate what "close" is. Including degrees of freedom benefits the reader in assessing the power of the tests used.

Summary statistics tables were reported in 52%

[3] Liu, Sun, and Kaiser did not elaborate on how it was determined that no significant impact on model results occurred from excluding a statistically insignificant variable.

of the articles surveyed (question 18). Surprisingly, many papers used means in interpreting results, e.g., computing elasticities, without reporting summary statistics and the reader was left to wonder what the summary statistics of the data were. Thirty percent of the articles surveyed did not report coefficients in a form that readers could easily interpret (question 19). Many of these articles were LDV models; however, a substantial number of the articles were OLS models (table 4). For example, Lansford and Jones reported results from an OLS regression model using Box-Cox transformed data. The data transformation was non-linear, making interpretation of the coefficients less than intuitive; yet, Lansford and Jones provided a subsequent table of interpretable marginal values.[4]

Questions 20 through 24 were posed to determine the interpretation of statistical and economic significance in the article. Seventy-three percent of the articles surveyed discussed both the statistical and economic importance of the variables (question 20), and in 35% of the articles the author emphasized statistical significance more than economic significance (question 21). As an example of discussing both statistical and economic significance, Holt quantified his results and stated, "Overall, the results indicated that the third-moment is both statistically significant and economically significant." (p. 251) In 66% of the applicable articles surveyed, the author discussed the size of the effects (question 22). Articles that used OLS, relative to LDV and SYS, tended to have more discussion on both statistical and economic significance (table 4). When an LDV model was used, the author was less likely to discuss the size of the effect (table 4). Similarly, a lower percentage of the articles reported the size of the coefficients when cross-sectional data were used relative to when time-series or panel data were used (table 5).

Eighty-six percent of articles reported results of coefficient signs consistent with the statistical significance of the coefficients (question 23). However, 14% of the articles were not consistent and this may be of concern in assessing the impacts of some of the variables. Authors were generally careful in using the word significance, so that statistical and economic significance were not confused (question 24). There were no discernable differences based on econometric method or data type for these two questions.

Only 38% of the articles surveyed used simulation to enhance results (question 25). This percentage is surprisingly low considering the applied nature of the agricultural economics profession.[5] Teasley, Bergstrom and Cardell estimated willingness to pay for public area recreation and used simulation to explicitly show the impact different fees would have on annual revenue. Perhaps simulation was not used in some studies because it would have indicated uninteresting results, it would not have added much of substance to the paper, or authors have not considered the potential merits of such an experiment.

Discussion

This study reviewed articles using regression analysis published in the *Journal of Agricultural and Resource Economics* from 1994 to 1998 to determine agricultural economists' effectiveness in reporting and conveying research procedures and results. Twenty-five questions were asked of each article. In addition to questions addressing the differentiation between statistical and economic significance, questions addressed use of data, estimation procedures, level of confidence indicated by the use of asterisks, and goodness of fit measures used by the author.

Economics is a social science relying heavily on data collection and analytical methods to validate theoretical hypotheses or determine generalities from observed data. Statistical versus economic significance, and the implications of each, has been even more of an issue for agricultural economists due to the applied nature of their research relative to the economics profession. In presenting the procedures for developing tests of statistical significance in regression analysis, Wiegmann (*Journal of Farm Economics* 1954) concluded the "Interpretation of the Results" section with

> It should be mentioned that a difference that is statistically significant is not the same as an economically important difference. Whether a difference is economically important depends on criteria other than statistics. The tests could be used, however, to test null hypotheses such as HO: $X_1 - X_2 = K$ where K is some amount great enough to make a statistical difference between the means, regression coefficients or other measures also important economically. (p. 639)

[4] Wilde and Ranney discussed results of models not reported in the study, but available from the authors. Though different from interpretation difficulties, this made evaluating results a bit confusing.

[5] Simulation may not be applicable to all economic research. As noted by one reviewer, simulating the quantity effects from a demand system may overstate the impact when supply is upward sloping.

Wiegmann's discussion was on the limitations and implications of statistical tests using OLS regression. During Wiegmann's era, which coincided with the beginning of reporting econometric methods (Debertin and Pagoulatos), a lack of statistical tools allowed for economic significance to be the focus. However, since Wiegmann's era, advances in analytical techniques have occurred and statistical methods to improve the quality and quantity of information contained in the data have been developed.[6] In the publishing game, what impact has uncovering statistical significance had on identifying economic significance? Are tests of statistical significance used to defend results? In the words of Popper, "whenever we try to propose a solution to a problem, we ought to try as hard as we can to overthrow our solution rather than defend it." (p. 7) Possibly, more emphasis should be placed on testing for model fragility than on tests of statistical significance.

Comparing McCloskey and Ziliak, who concluded that over three-fourths of their surveyed articles misused the test of statistical significance, to results here, agricultural economists are better than general economists at differentiating tests of statistical significance from tests of substantive difference. Viewing the question of differentiating statistical significance from social significance, Neyman and Pearson postulated, "Is it more serious to convict an innocent man or acquit a guilty? That will depend on the consequences of the error . . ." (p. 296, cited in McCloskey and Ziliak p. 97). Perhaps the costs (professionally) to agricultural economists of not considering economic implications are greater because of the applied nature of our research.

Based on the authors' experiences of surveying articles for this study, we have several suggestions on how to better express reporting of results and how to better separate statistical from economic significance. First, clearly define the dependent variable—preferably in the results table as well as written text. Second, when applicable, report parameter estimates in an interpretable form either in the results table or in a subsequent table. Third, report summary statistics. What might seem redundant to an author could be essential to a reader. Fourth, report degrees of freedom conspicuously in the results table. Fifth, report if statistically insignificant variables have been dropped. Lastly, weigh economic importance aside from statistical significance—use simulation as necessary or when useful to express economic significance. Another consideration, not implicitly examined in this study, is the usefulness of the author to report limitations and possible extensions of the research—no one knows better than the author.

This study was not intended to compromise any one person's or group's past or future research efforts. Published research, even if it has weaknesses, is still superior to unpublished work (even without weaknesses). Because, as Brorsen argued (citing Adams), ". . . too much importance is given to the number of publications with only limited emphasis on the quality, but research that is never communicated to others is indeed of little value." (p. 315) The current paper sought to improve communications between authors and readers.

[6] Evidence of this is the study by Debertin and Pagoulatos in which they surveyed articles published in the *American Journal of Agricultural Economics* between 1919 and 1990. From 1950 to 1959, 17% of the articles published used either single or simultaneous estimation, and from 1981 to 1990, 54% of the articles published used either single or simultaneous estimation.

References

Adams, D.W. 1985. "Assessing the Usefulness of Publications by Agricultural Economists Through Citations." The Ohio State University, Economics and Sociology Occasional Paper No. 1215, October 18.

Blaug, M. 1992. *The Methodology of Economics*. New York: Cambridge University Press.

Brorsen, B.W. 1987. "Observations on the Journal Publication Process." *North Central Journal of Agricultural Economics* 9:315–21. July.

Debertin, D.L., and A. Pagoulatos. 1992. "Research in Agricultural Economics 1919–1990: Seventy-Two Years of Change." *Review of Agricultural Economics* 14:1–22.

Griliches, A. 1986. "Productivity, R&D, and Basic Research at the Firm Level in the 1970s." *American Economic Review* 76:141–54. December.

Helmberger, P., and Y. Chen. 1994. "Economic Effects of U.S. Dairy Programs." *Journal of Agricultural and Resource Economics* 19:225–38. December.

Holt, M. 1994. "Price-Band Stabilization Programs and Risk: An Application to the U.S. Corn Market." *Journal of Agricultural and Resource Economics* 19:239–54. December.

Hurd, B.H. 1994. "Yield Response and Production Risk: An Analysis of Integrated Pest Management In Cotton." *Journal of Agricultural and Resource Economics* 19:313–26. December.

Kenkel, P.L., and P.E. Norris. 1995. "Agricultural Producers' Willingness to Pay for Real-Time Mesoscale Weather Information," *Journal of Agricultural and Resource Economics* 20:356–72. December.

Ladd, G.W. 1991. "Thoughts on Building an Academic Career." *Western Journal of Agricultural Economics* 16:1–10. July.

Lansford, N.H., and L.L. Jones. 1995. "Recreational and Aesthetic Value of Water Using Hedonic Price Analysis."

Journal of Agricultural and Resource Economics 20:341–55. December.

Levins, R.S. 1992. "The Whimsical Science." *Review of Agricultural Economics* 14:139–51. January.

Liu, D.J., C. Sun, and H.M. Kaiser. 1995. "Market Conduct under Government Price Intervention in the U.S. Dairy Industry." *Journal of Agricultural and Resource Economics* 20:301–15. December.

McCloskey, D.N., and S.T. Ziliak. 1995. "The Standard Error of Regression." *Journal of Economic Literature* 97–114. March.

McCloskey, D.N. 1985. "The Loss Function Has Been Mislaid: The Rhetoric of Significance." *American Economic Review* 75:201–5. May.

Neyman, J., and E.S. Pearson. 1933. "On the Problem of the Most Efficient Tests of Statistical Hypothesis." *Philosophical Transaction of the Royal Society* A 231:289–37.

Pritchett, J.G., D.J. Liu, and H.M. Kaiser. 1998. "Optimal Choice of Generic Milk Advertising Expenditures by Media Outlet." *Journal of Agricultural and Resource Economics* 23:155–69. July.

Popper, K. 1965. *The Logic of Scientific Discovery.* New York: Harper Torchbooks.

Ramezani, C.A. "Determinants of Nutrient Demand: A Non-parametric Analysis." 1995. *Journal of Agricultural and Resource Economics* 20:166–77.

Teasley, R.J., J.C. Bergstrom, and H.K. Cordell. 1994. "Revenue-Capture Potential of Public Area Recreation." *Journal of Agricultural and Resource Economics* 19:89–101. July.

Tomek, W.G. 1993. *American Journal of Agricultural Economics* 75, 75th Anniversary Issue:6–14. October.

Tweeten, L. 1983. "Hypothesis Testing in Economic Science." *American Journal of Agricultural Economics* 65:548–52. August.

Wiegmann, F.H. 1954. "Use of Regressions and Test Procedures." *Journal of Farm Economics* 36:633–40. November.

Wald, A. 1939. "Contribution to the Theory of Statistical Estimation of Testing Hypotheses." *Annals of Mathematical Statistics* 10:299–326. December.

Wallace, T.D. 1977. "Pretest Estimation in Regression: A Survey." *American Journal of Agricultural Economics* 59: 431–43. August.

Wilde, P., and C. Ranney. 1996. "The Distinct Impact of Food Stamps on Food Spending." *Journal of Agricultural and Resource Economics* 21:174–85. July.